ORIGINAL PAPER

Improving the robustness of biomass functions: from empirical to functional approaches

**Ablo Paul Igor Hounzandji · Mathieu Jonard ·
Claude Nys · Laurent Saint-André · Quentin Ponette**

Abstract
· *Key message* We developed precise, consistent, generic, and robust biomass equations for seven aboveground tree components of sessile and pedunculate oaks. These equations can be used to accurately estimate carbon stocks and fluxes in and out of the forest.
· *Context* Large uncertainties still persist when using existing biomass equations for larger scale applications.
· *Aims* The objective of this study was to test two contrasting modeling approaches to obtain biomass estimates of various components (stem, stem wood, stem bark, crown, and three branch categories) for *Quercus petraea* and *Quercus robur* and to compare them in terms of predictive capacity, genericity, consistency, and robustness.
· *Methods* All models were calibrated on a total of 117 oak trees sampled over a wide range of sites and stands and further tested on an independent data set of 33 trees. The "empirical" approach consisted in declining a common allometric equation based on two variables (diameter at breast height and total height) into all its possible forms and selecting the final model on purely statistical performances; the "structural" method was based on the fitting of a priori dedicated model forms for each component to allow a clear interpretation of the model parameters.
· *Results* For the stem components, both approaches resulted in similar statistical performances despite difference in model forms and number of parameters. Although equally performant on the validation data set for the total crown, only the structural model gave satisfactory results when applied to the independent data set. Both approaches failed to accurately predict the branch fractions on the validation data set.
· *Conclusion* Using physically based model forms increased the robustness of the biomass equations.

Contribution of the co-authors API H carried out the data analysis and contributed to the text; M J contributed to the supervision of the work and to the text; C N provided most of the raw biomass data; L SA contributed to the development of the modeling framework; Q P obtained the grant, supervised the work, and contributed to the text.

A. P. I. Hounzandji · M. Jonard · Q. Ponette (✉)
Earth and Life Institute, Environmental Sciences, UCL, Croix du Sud, 2, box L7.05.09, 1348 Louvain-la-Neuve, Belgium
e-mail: quentin.ponette@uclouvain.be

A. P. I. Hounzandji
e-mail: ablo.hounzandji@uclouvain.be

M. Jonard
e-mail: mathieu.jonard@uclouvain.be

C. Nys · L. Saint-André
UR1138, Unité Biogéochimie des Ecosystèmes Forestiers (BEF), Centre INRA de Nancy, INRA, 54280 Champenoux, France

C. Nys
e-mail: nys@nancy.inra.fr

L. Saint-André
e-mail: st-andre@nancy.inra.fr

L. Saint-André
CIRAD, UMR ECO&SOLS, 34000 Montpellier, France

1 Introduction

Forests have been pointed out as key components of the global carbon cycle since these ecosystems store large reservoirs of carbon (more than two times the amount of carbon in the atmosphere) and absorb about 30 % of all CO_2 emissions (Sabine et al. 2004). In the global change context, forestry activities may contribute to mitigating carbon emissions in four ways: (i) by increasing reforestation, (ii) by reducing

deforestation, (iii) by favoring practices that increase carbon storage in existing forests, and (iv) by expanding the use of wood products (as material or energy) in order to limit the consumption of fossil fuels and to favor carbon storage outside the forest (Canadell and Raupach 2008). Regarding the European forest, mainly the last two mitigation strategies are possible since we cannot expect drastic change in forest cover (FAO 2010). As both strategies can lead to contradictory guidelines, integrated approaches must be developed to find the optimal way of managing forest carbon (McKechnie et al. 2010). Such approaches require accurate estimates of carbon fluxes into and out of the forest as well as of carbon stocks within the forest and in wood products (Vanclay and Skovsgaard 1997; Ketterings et al. 2001; Fortin et al. 2012). These forest carbon estimates are also crucial for carbon credit trading and national reporting of greenhouse gas inventories to the United Nations Framework Convention for Climate Change.

The carbon amount of individual trees can be estimated based on biomass equations relating variables that are commonly measured during forest inventories (e.g., diameter at breast height and total tree height) to dry mass of tree components (fine roots, coarse roots, trunk, branches, foliage). Alternatively, biomass expansion factors can also be used (Lehtonen et al. 2004; Longuetaud et al. 2013). Since the early 1970s, many biomass equations have been established for a wide variety of tree species in different regions of the world (Ter-Mikaelian and Korzukhin 1997; Nelson et al. 1999; Lambert et al. 2005; Saint-André et al. 2005; Wutzler et al. 2008). Nevertheless, the users of these biomass equations are generally facing several problems:

(i)   The growth conditions (soil, climate) and the forestry practices of the sites in which biomass equations are used often differ from those of the sites that served to develop these equations (Ter-Mikaelian and Korzukhin 1997).

(ii) Various biomass equations are available for a same tree species but are not always directly comparable since they may differ by the mathematical form of the models (linear or nonlinear), the way the error term is considered (additive or multiplicative error), the way the variance is dealt with (constant or not), the transformation of the data (e.g., logarithmic vs square-root), and the selected predictor variables (Parresol 1999).

(iii) Biomass equations of the different tree components were often adjusted independently, and additivity is not necessarily assured (estimated biomass of the tree components was not constrained to equal estimated total biomass).

(iv) Information on model performances are generally limited and not evaluated on independent data sets (Prisley and Mortimer 2004).

To overcome these issues, new approaches have been developed recently. While biomass studies often relied on limited numbers of sites and trees in the past, some biomass data from various sources were recently pooled together to obtain large data sets, which therefore broadened the validity range of the biomass equations established based on these data (Wutzler et al. 2008). In addition, when a lot of data are available, a subset can be devoted to validation in order to independently evaluate model performances (Saint-André et al. 2005; Wutzler et al. 2008; Genet et al. 2011). The use of covariate models is another way to improve the validity range of biomass equations; this approach consists in fitting biomass equations site by site (or stand by stand) and to relate equation parameters to site/stand variables (Saint-André et al. 2005; António et al. 2007; Genet et al. 2011). Finally, the problem of tree component additivity is now commonly solved by using joint-generalized least square regression, also called seemingly unrelated regression (Parresol 2001; Lambert et al. 2005).

In this study, we aimed at developing biomass equations for several tree components of sessile and pedunculate oaks (*Quercus petraea* Matt. Liebl. and *Quercus robur* L.). Although sessile and pedunculate oaks are dominant broad-leaved species in Europe (Zanetto et al. 1994), they have been little studied regarding biomass; most of the existing biomass equations were developed based on a limited number of trees and concern mainly sessile oak (André and Ponette 2003; Cienciala et al. 2008; André et al. 2010; Suchomel et al. 2012). For this study, we used a large amount of data obtained from 150 trees originating from 16 stands (18 to 182 years old) distributed over seven French and Belgian sites.

The objective of our study was to develop biomass equations using two contrasted approaches and to compare them with existing equations by considering the predictive capacity, the genericity, the consistency, and the robustness of the models. Using a well-defined model selection procedure, the first approach called "empirical" consisted in integrating stand effects in a classical allometric equation based on two variables (diameter at breast height and total tree height) and applied indifferently to all tree components. The advantage of such an approach is to broaden the range of situations where the model can be used. The second approach called "structural" was less systematic regarding the model selection since the objective was to develop physically meaningful equations based on variables that specifically characterize the considered tree component. In this approach, the selected model form depends therefore on the tree component. For the trunk, we described the biomass as the product between the volume and the density, where the volume was determined based on the trunk height and diameter and on a form coefficient. For the branches, we considered an estimate of the diameter at crown base as well as the crown length. Our hypothesis was that similar model performances will be obtained by both approaches based on the calibration data set (equivalent prediction accuracy) but that these performances

will be higher for the structural approach when considering an independent validation data set (higher robustness for the structural approach). Depending on their performances, models from the structural approach may consist in a step forward toward the development of more generic models. In addition, the values of physically meaningful parameters of the structural models (form coefficient, wood density) were compared with measurements made in other studies.

2 Material and methods

2.1 Site description and aboveground biomass measurements

2.1.1 Calibration data set

A total of 117 oak trees (62 sessile oak (*Q. petraea* Liebl.) and 55 pedunculate oak trees (*Q. robur* L.) were sampled in five French sites characterized by elevations ranging from 116 to 320 m, mean annual temperatures varying from 7 to 11 °C, and total annual precipitation varying from 600 to 1,100 mm (Table 1). The characteristics of the selected stands and trees are presented in Table 2. The sampled trees spanned a wide range of ages (18 to 150 years), diameters at 1.3-m height ($d=$ 9–73 cm), and total heights ($ht=$8–32 m); they originated from even-aged pure oak (Potées, Perche-Trappe, Blois, and Bercé) or mixed-species stands dominated by beech (Fougères), as well as from fertilized (Ca, Mg) and unfertilized sites. Thus, our sample covered a wide range of site conditions and stand characteristics.

The biomass measurements can be summarized as follows: (i) stand inventory and selection of a sample of trees covering the range of tree diameters at breast height, (ii) felling of the selected trees in the resting period and tree measurements (total height, diameter at the felling section, at breast height, and at every 1 m along the stem), (iii) separation of both trunk (defined as the main axis from tree base to the apex) and branches into three-diameter classes (0–4, 4–7, and >7 cm), resulting in a total of maximum six woody components per tree, (iv) total fresh weight determination of each component in the field, and (v) taking of subsamples for determination of fresh to dry weight ratios (65 °C; all woody components) and wood to bark weight ratios (trunk diameter >7 cm). For a limited number of trees, the lower part of the trunk could not be weighted in the field; in those cases, the corresponding biomass was obtained from the green volume and the basic wood density estimated from a subsample of tree boles of comparable size. Additional information on biomass measurements is given by Ranger et al. (1995).

2.1.2 Validation data set

The validation data set was made up of 33 sessile oaks from two unfertilized Belgian sites (Tables 1 and 2); one stand had an almost regular structure (site 6), while the other was more irregular (site 7) as it originated from a former coppice with standards under conversion. The sampled trees (Table 2) were aged from 38 to 182 years with diameters ranging from 16 to 66 cm and total heights ranging from 16 to 27 m. The biomass measurements (André and Ponette 2003; André et al. 2010) were comparable to those used for the calibration data set, except for the division of the trees into compartments. In this case, the trunk was defined as the main axis from tree base to the so-called Delevoy height (*h*—height at which trunk diameter is half the diameter at breast height). The parts of the trees above the Delevoy height together with the branches attached to the trunk were considered as crown.

2.2 Biomass components and model development

2.2.1 Biomass components

The final biomass components considered in this study are the stem (total stem, stem bark, stem wood), the crown, the small branches, the medium branches, and the coarse branches. To allow comparison between the calibration and the validation data sets, the stem was similarly defined as being the main axis extending from the felling section to the Delevoy height,

Table 1 Site characteristics (calibration and validation data sets)

Type of data	Site number	Site name	Latitude	Longitude	Elevation (m)	T (°C)	P (mm)
Calibration	1	Fougères	48° 23′ 07″ N	1° 10′ 50″ O	155–170	11.1	990
	2	Potées	49° 51′ 32″ N	4° 29′ 39″ E	320	7.3	1,100
	3	Perche-Trappe	48° 40′ 39″ N	7° 04′ 04″ E	175–320	10.1	750
	4	Blois	47° 33′ 03″ N	1° 15′ 04″ E	116	10.5	690
	5	Bercé	47° 48′ 05″ N	0° 23′ 52″ E	179	11.1	678
Validation	6	Baileux	50° 01′ 00″ N	4° 24′ 00″ E	300	8	1,044
	7	Chimay	50° 06′ 46″ N	4° 17′ 05″ E	250	8.5	1,130

T annual mean temperature, *P* annual total precipitation

Table 2 Characteristics of the selected stands and trees (calibration and validation data sets)

	Selected stands						Selected trees				
Type of data	Age (year)	Site number	Oak species	Fertilization	G (m^2 ha^{-1})	$Hdom$ (m)	N	D (cm)	Ht (m)	W_{stem} (kg)	W_{crown} (kg)
Calibration	18	2	Qr	U	12.6	12.9	8	17.4	12.9	48.3	43.6
	18	2	Qr	F	13.1	12.7	7	17.9	12.7	55.9	40.5
	18	3	Qp	F	17.1	7.7	2	14.1	8.3	32.2	29.4
	25	3	Qr, Qp	U	26.2	14.2	7	17.6	13.2	80.5	35.1
	35	4	Qp	U	32.6	13.5	6	17.6	13.3	91.1	46.4
	35	4	Qp	F	32.8	13.2	5	20.2	13.4	117.1	50.8
	45	3	Qr, Qp	U	23.5	18.0	12	22.7	17.1	190.5	74.7
	88	1	Qr	U	NA	NA	2	30.8	25.4	489.6	118.9
	88	5	Qp	U	46.7	26.9	12	31.8	26.3	593.4	187.7
	88	5	Qp	F	45.9	27.1	12	31.7	25.9	542.2	193.7
	90	1	Qr	U	NA	NA	9	28.5	21.3	348.7	169.1
	100	3	Qr, Qp	U	24.0	27.7	12	37.3	27.7	742.5	269.0
	135	3	Qr, Qp	U	28.7	30.5	11	48.4	29.6	1,165.0	534.2
	150	3	Qr, Qp	U	38.3	30.2	12	56.3	29.5	1,521.5	825.6
Validation	62–129	6	Qp	U	23.7	24	15	37.1	22.7	764.3	389.7
	38–182	7	Qp	U	16.8	NA	18	43.4	19.6	864.7	389.5

Qr Quercus robur, *Qp Quercus petraea*, *U* unfertilized, *F* fertilized, *G* stand basal area, *Hdom* dominant height, *N* number of sampled trees, *D* mean diameter at 1.30 m, *Ht* mean total height, *W* mean biomass, *NA* data not available

while the remaining living woody parts (excluding epicormic branches) were considered as the crown and further subdivided into three categories: small ($d=0–4$ cm), medium ($d=4–7$ cm), and coarse ($d>7$ cm) branches. For the calibration data set, the stem biomass was calculated by linear interpolation between the corresponding successive trunk sections, while that part of the trunk located beyond the Delevoy height was added to the corresponding branch diameter class(es).

2.2.2 Model development

Empirical approach

In this purely statistical approach, we started from the following reference equation (Genet et al. 2011):

$$Y = a + \beta(X)^\gamma + \varepsilon \tag{1}$$

Y is the dry biomass (kg) of a given tree compartment, X is either (d^2h) for the stem components or (d^2ht) for the crown components, with h being Delevoy height (m) and ht being total tree height (m), and α, β, and γ are the model parameters to be estimated.

Taking into account the specific cases where $\alpha=0$ and/or $\gamma=1$, this resulted in a set of four basic equations. These equations were fitted considering all stands together (global fitting; models 1 to 4 in Table 3) as well as stand by stand for each basic equation (so-called local fitting; models 1.1 to 1.7, 2.1 to 2.3, 3.1 to 3.3, and 4.1 in Table 3). In stands where liming or fertilization was applied, the fertilized and unfertilized parts of a same stand were considered together for the analysis. In addition, two stands with only two sampled trees were added to other stands of the same age.

The resulting 18 models (Table 3) were adjusted for each compartment, and the selection of the final model was based on the following four (i to iv) steps:

(i) We first used the Akaike information criterion (AIC) and Bayesian information criterion (BIC) criteria (Schwarz 1978; Sakamoto et al. 1986), and the best model was that with the lowest values for both criteria. When results from AIC and BIC criteria did not match, we performed a likelihood ratio test (LRT); in the case of a significant LRT, we selected the model with more parameters, otherwise we retained the simpler model.

(ii) When the model selected in the previous step was a global model, the selection process ended. By contrast, when a significant stand effect was detected, a covariate was further tested for possible inclusion by modeling the relationship between the stand-dependent parameter(s) of the model and stand age.

(iii) In the case of a significant covariate effect, the latter was included in the model, and the resulting so-called covariate model was refitted and compared to the corresponding global model using AIC and BIC criteria as in step (i).

Table 3 Deterministic terms of basic and local models (empirical approach)

Basic	$Y=a+\beta X^{\gamma}$	(1)[a]	$Y=\beta X^{\gamma}$	(2)	$Y=a+\beta X$	(3)	$Y=\beta X$	(4)
Local	$Y_s=a_s+\beta_s X^{\gamma s}$	(1.1)	$Y_s=\beta_s X^{\gamma s}$	(2.1)	$Y_s=a_s+\beta_s X$	(3.1)	$Y_s=\beta_s X$	(4.1)
	$Y_s=a_s+\beta_s X^{\gamma}$	(1.2)	$Y_s=\beta_s X^{\gamma}$	(2.2)	$Y_s=a_s+\beta X$	(3.2)		
	$Y_s=a_s+\beta X^{\gamma s}$	(1.3)	$Y_s=\beta X^{\gamma s}$	(2.3)	$Y_s=a+\beta_s X$	(3.3)		
	$Y_s=a+\beta_s X^{\gamma s}$	(1.4)						
	$Y_s=a_s+\beta_s X^{\gamma}$	(1.5)						
	$Y_s=a+\beta_s X^{\gamma}$	(1.6)						
	$Y_s=a+\beta X^{\gamma s}$	(1.7)						

The subscript s refers to a parameter estimate for a specific stand (local fitting)

Y biomass; X predictor; α, β, γ model parameters; X $d^2 h$ for stem, stem wood, and stem bark biomass equations; X $d^2 ht$ for crown, small branches, medium branches, and coarse branches biomass equations, d diameter at breast height; h Delevoy height; ht total height

[a] Reference equation

Based on this test, the best model, either covariate or global, was retained for further processing (see Sect. 2.3).

(iv) When the observed relationship between the local parameters and stand age was not statistically significant in step (ii), the analysis resumed from step (i), starting from the second best model and so on.

Structural approach
The so-called structural approach was based on an a priori selection of physically and/or biologically meaningful models for both the trunk and crown components.

Trunk components For the trunk components, we considered the stem, stem wood, and stem bark compartments. Since the biomass of any component is the product of its volume by its basic density, we started by modeling the outside-bark stem volume (V_{stem}, m³), using Eq. 5 where b_{stem} can be interpreted as the outside-bark form factor. This model is known as combined variable or constant form factor equation (Spurr 1952; Husch et al. 1982).

$$V_{stem} = b_{stem}(d^2 h) + \varepsilon_{stem} \quad (5)$$

As for the empirical approach, a local fitting of the model was carried out, and in the case of a significant stand effect, a covariate model was also tested.

In a second step, the inside-bark stem volume or stem wood volume ($V_{stem\ wood}$, m³) was modeled by Eq. 6 where du (m) is the under-bark diameter at breast height and h is defined as before.

$$V_{stemwood} = b_{stemwood}(du^2 h) + \varepsilon_{wood} \quad (6)$$

As bark thickness was not measured on the sampled trees, du was estimated from d, using the stem bark thickness model proposed by Dagnelie et al. (1999).

By hypothesizing the similarity between the outside- and inside-bark form factors ($b_{stem}=b_{stem\ wood}$), we could rewrite Eq. 6 as

$$V_{stemwood} = b_{stem}(du^2 h) + \varepsilon_{wood} \quad (7)$$

The biomass of the stem (Y_{stem}, kg) and stem wood ($Y_{stem\ wood}$, kg) compartments was then obtained by multiplying the corresponding volumes and basic densities, according to

$$Y_{stem} = \rho_{stem} b_{stem}(d^2 h) + \varepsilon_{stem} \quad (8)$$

$$Y_{stemwood} = \rho_{stemwood} b_{stem}(du^2 h) + \varepsilon_{stemwood} \quad (9)$$

In those equations, ρ_{stem} and $\rho_{stem\ wood}$ are the basic stem (wood+bark) and wood densities, respectively. As for the volume model, local and covariate models were fitted to test for a possible stand age effect on ρ_{stem} and $\rho_{stem\ wood}$.

The stem bark biomass ($Y_{stem\ bark}$, kg) was estimated as the difference between the stem and stem wood biomasses, which conferred a recursive character to the trunk biomass models (Dagnelie 1975):

$$Y_{stembark} = b_{stem}\{\rho_{stemwood}(d^2 h) - \rho_{stemwood}(du^2 h)\} + \varepsilon_{stembark} \quad (10)$$

Crown components For the crown components, we first modeled the total crown biomass and then derived the biomass of the branch categories from biomass ratio models.

The total crown biomass was modeled using the same strategy as that developed for the empirical approach but testing a wider range of model forms (linear, exponential, power, polynomial) and tree predictors (diameter at 1.3 m, d; total tree height, ht; the ratio ht/d, as an index of past competition; $d^2 ht$; $d^2 L$, where L=crown length=$ht-h$). To reduce the

number of possible combinations, we only tested models including one tree predictor plus a possible age effect introduced as a covariate.

The selected model was a simple linear model with d^2L as independent variable (Eq. 11):

$$Y_{crown} = \omega_0 + \omega_1(d^2L) + \varepsilon_{crown} \qquad (11)$$

Y_{crown} (kg) is the total crown biomass of a tree, and ω_0 and ω_1 are the model coefficients to be estimated.

For the branch categories, we developed a specific biomass ratio model for the small branches $\hat{R_1}$ (Eq. 12) and coarse branches $\hat{R_3}$ (Eq. 13), according to

$$\widehat{R}_1 = \theta_0 + e^{-\theta_1(Y_{crown})^{\theta_2}} + \varepsilon \qquad (12)$$

$$\widehat{R}_3 = \frac{\theta_3 Y_{crown}}{\theta_4 + Y_{crown}} + \varepsilon \qquad (13)$$

Y_{crown} is the crown biomass estimated by Eq. 11, and θ_1, θ_2, θ_3, and θ_4 are the model parameters to be estimated.

The medium branch ratio was estimated by difference $\hat{R_2}$ (Eq. 14), taking into account the ratio unit constraint ($\sum$ratios=1):

$$\widehat{R}_2 = 1 - \left(\widehat{R}_1 + \widehat{R}_3\right) + \varepsilon \qquad (14)$$

2.3 Model fitting

For each approach, the selected biomass models were simultaneously adjusted to satisfy the biomass compartment additivity constraint (Parresol 1999, 2001; Návar et al. 2002; Carvalho and Parresol 2003; Bi et al. 2004; Genet et al. 2011). Seemingly unrelated regression method was used to solve the simultaneous equation system through the PROC MODEL procedure in SAS (SAS Institute Inc. 2011), and the parameter values were derived using the maximum likelihood method. In the empirical approach, we experienced some difficulties in fitting the global equation system (lack of convergence, biased estimates). We therefore performed an adjustment in three stages: (i) additivity of stem wood and stem bark biomass with respect to the stem, (ii) additivity of the branch categories biomass with respect to the crown, and (iii) additivity of the stem and crown biomass with respect to the total tree biomass. In the structural approach, the recursive nature of the biomass models already satisfied the additivity constraint. The system of equations under consideration consisted in stem volume (Eq. 5), stem biomass (Eq. 8), stem bark biomass (Eq. 10), and crown biomass

(Eq. 11) models. As the stem bark model (Eq. 10) contained all the parameters of the stem wood model (Eq. 9), we removed the stem wood model from the system in order to give more weight to the stem bark model.

To correct for heteroscedasticity, the variance of the residuals was explicitly modeled, using either a power function for the biomass and volume models or an exponential function for the biomass ratio models (Carroll et al. 1988; Pinheiro and Bates 2000; Huet et al. 2004).

2.4 Final evaluation of the selected model

2.4.1 Calibration data set

The statistical performances of the simultaneously adjusted models were first examined on the calibration data set through the following quantitative indicators: the root mean square error in percent (%RMSE) (Arevalo et al. 2007), the modeling efficiency (EF) (Loague and Green 1991; Mayer and Butler 1993), and the simultaneous F-test of bias (slope=1 and intercept=0 for the regression between observed and predicted values) (Dent and Blackie 1979; Mayer et al. 1994). Because the simultaneous F-test may be quite sensitive (Thornton and Hansen 1996; Analla 1998), we considered a detected significant bias to be still acceptable if the confidence interval of the intercept contained zero and that of the slope contained 1. Finally, the normality of the standardized residuals was checked using a *qqplot*, and their equivariance and independence were checked by graphic representation of the couples (residuals, predicted values).

2.4.2 Validation data set

These criteria were also applied to the validation data set, using the same models as above as well as two additional models available from the literature (Cienciala et al. 2008; Suchomel et al. 2012). Because the definition of the compartments (stem, stem wood, stem bark, total crown) did not match exactly between these studies and ours, the models taken from the literature were first fitted on our calibration data set (Tables 5 and 6). Among the three functions developed by Cienciala et al. (2008), only those two that incorporated either d or d and h could be tested, and the best model based on the AIC, BIC, or *LRT* criteria was retained for comparison with our equations. In order to compare all models on a strictly similar basis, the equations from Cienciala et al. (2008) were not linearized by log-transformation prior to fitting as done in the original paper.

3 Results

3.1 Empirical approach

3.1.1 Calibration data set

The AIC and BIC values resulting from the fitting of the 18 models (Table 3) to each of the seven compartments (122 adjustments out of a total of 126 possible cases) are reported in Table 4. One stem (model 1.1 for stem bark) and three branch (model 1.1 for the medium branches; models 2.1 and 3.1 for the coarse branches) models could not be obtained due to a lack of convergence during the fitting procedure. For all the compartments with the exception of the coarse branches, an appropriate covariate model could be successfully derived from the corresponding local model; in all cases, these covariate models had comparable AIC and BIC values than their local counterparts and consistently lower values than the corresponding global models. For the coarse branches, the modeling process ended with the selection of a global model (model 2).

All those models were then simultaneously adjusted using seemingly unrelated regression to ensure additivity, and their resulting parameter values as well as statistical performances are presented in Tables 5 and 6, respectively.

For the three stem components (total stem, stem wood, and stem bark), all models derived from the same base model (model 1 in Table 3) and led to unbiased (non significant F-test) and high precision (%RMSE 11.69–25.73; EF 0.90–0.99) estimates. All models were characterized by a fixed β value and an age-dependent γ parameter, while α was age-dependent only for the total stem and stem wood components.

The final crown models derived from three different base models (small branches, model 1; medium and coarse branches, model 2; total crown, model 3). The precisions were lower (%RMSE 40.63–69.27; EF 0.73–0.87) than for the stem components, and although significant, the bias proved to be acceptable. For the three local models, the covariate-related stand age effect was associated either to the intercept (total crown and small branches; additive

Table 4 Information criterion statistics (*AIC* and *BIC*) for the whole set of models tested in the empirical approach

Model[a]	Stem		Stem wood		Stem bark		Crown		Small branches		Medium branches		Coarse branches	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
(1)	1,175.45	1,189.26	1,148.92	1,162.73	810.26	824.07	1,251.27	1,265.08	986.81	1,000.62	925.72	939.54	1,105.59	1,119.40
(1.1)	1,132.73	1,212.84	1,112.26	1,192.36	NC	NC	1,222.37	1,302.47	983.99	1,064.10	NC	NC	1,087.57	1,167.67
(1.2)	1,129.79	1,187.79	1,111.07	1,169.08	755.38	813.38	1,213.70	1,271.71	971.11	1,029.11	897.69	955.69	*1,083.12*	*1,141.52*
(1.3)	***1,120.15***	***1,178.15***	***1,102.02***	***1,160.03***	750.77	808.78	1,221.30	1,279.31	970.47	1,028.47	897.85	955.86	1,086.32	1,144.32
(1.4)	1,131.17	1,189.18	1,112.26	1,170.27	753.12	811.13	1,213.61	1,271.61	970.35	1,028.35	896.23	954.24	1,084.15	1,142.16
(1.5)	1,170.63	1,206.54	1,145.52	1,181.43	793.03	828.93	1,210.78	1,246.68	***961.44***	***997.35***	891.75	927.66	1,090.95	1,126.86
(1.6)	1,126.47	1,162.37	1,106.63	1,142.54	753.01	788.92	1,215.55	1,251.45	966.59	1,002.50	894.07	929.98	1,086.80	1,122.70
(1.7)	1,127.70	1,163.60	1,109.33	1,145.23	***745.78***	***781.69***	1,244.90	1,280.81	981.93	1,017.84	924.67	960.58	1,091.88	1,127.79
(2)	1,174.17	1,185.22	1,147.95	1,159.00	807.42	818.47	1,261.30	1,272.34	989.90	1,000.95	928.19	939.24	**1,103.03**	**1,114.08**
(2.1)	1,135.23	1,190.47	1,115.49	1,170.74	756.24	811.48	1,212.43	1,267.67	967.75	1,022.99	893.61	948.85	NC	NC
(2.2)	1,148.18	1,181.32	1,123.60	1,156.75	779.71	812.86	1,230.63	1,263.78	966.12	999.27	***891.84***	***924.99***	1,095.47	1,128.61
(2.3)	1,162.68	1,195.83	1,138.13	1,171.28	791.03	824.17	1,244.18	1,277.32	980.37	1,013.51	925.35	958.50	1,095.96	1,129.11
(3)	1,175.00	1,186.05	1,146.81	1,157.86	822.81	833.86	1,304.78	1,315.83	1,000.92	1,011.97	965.32	976.37	1,216.79	1,227.84
(3.1)	1,127.87	1,183.11	1,109.09	1,164.33	754.49	809.74	1,220.03	1,275.27	968.30	1,023.55	894.85	950.09	NC	NC
(3.2)	1,174.48	1,207.63	1,154.21	1,187.36	789.80	822.94	***1,208.89***	***1,242.03***	980.76	1,013.91	920.15	953.30	1,214.39	1,247.53
(3.3)	1,130.20	1,163.35	1,113.14	1,146.29	752.43	785.57	1,297.27	1,330.41	979.06	1,012.20	909.54	942.68	1,249.06	1,282.20
(4)	1,173.02	1,181.31	1,147.98	1,156.27	873.85	882.14	1,261.96	1,270.24	1,069.37	1,077.66	984.05	992.34	1,165.69	1,173.98
(4.1)	1,146.22	1,176.60	1,121.87	1,152.25	808.46	838.84	1,247.17	1,277.55	966.20	996.58	896.12	926.50	1,139.91	1,170.29
CM	1,114.73	1,134.07	1,093.15	1,112.49	763.07	779.65	1,237.25	1,251.06	964.47	981.04	897.78	911.59	NA	NA

Best candidate models based on the AIC and BIC values or on the Likelihood ratio test are in italics. Final global or local model retained at the end of the four-step procedure (see text for details) is in bold

NC not converged, *NA* not available, *CM* covariate model, *AIC* Akaike information criterion, *BIC* Bayesian information criterion

[a] Model numbers refer to Table 3

Table 5 Parameter values of the selected models

Approach	Type		Model	Deterministic term	Error term (ε)
Empirical[a]	Biomass	Stem	(1.3)	$Y_{stem}=(39.25-0.76age)+331.53\ (d^2h)^{(age/(-10.49+1.16age))}$	$\varepsilon=26.87(d^2h)^{0.98}$
		Stem wood	(1.3)	$Y_{wood}=(32.28-0.69age)+298.88\ (d^2h)^{(age/\ (-9.78\ +\ 1.14age))}$	$\varepsilon=24.39(d^2h)^{1.04}$
		Stem bark	(1.7)	$Y_{bark}=5.29+28.61\ (d^2h)^{(age/\ (-15.03\ +\ 1.32age))}$	$\varepsilon=6.19(d^2h)^{0.80}$
		Crown	(3.2)	$Y_{crown}=(29.99-0.59age)+63.20d^2ht$	$\varepsilon=76.78(d^2ht)^{0.52}$
		Small branches	(1.5)	$Y_{sb}=(7.57-0.37age)+41.62\ (d^2ht)^{0.60}$	$\varepsilon=10.68(d^2ht)^{0.54}$
		Medium branches	(2.2)	$Y_{mb}=(29.97-0.15age)\ (d^2ht)^{0.95}$	$\varepsilon=7.89(d^2ht)^{0.64}$
		Coarse branches	(2)	$Y_{cb}=19.32\ (d^2ht)^{1.52}$	$\varepsilon=14.48(d^2ht)^{1.22}$
Structural	Volume	Stem	(5)	$V=0.52\ (d^2h)$	$\varepsilon=0.04(d^2h)^{0.98}$
	Biomass	Stem	(8)	$Y_{stem}=562.17\ (V)$	$\varepsilon=35.47(d^2h)^{0.70}$
		Stem wood	(9)	$Y_{wood}=594.92\times0.52\ (du^2h)$	$\varepsilon=35.57(d^2h)^{0.73}$
		Stem bark	(10)	$Y_{bark}=Y_{stem}-Y_{wood}$	$\varepsilon=7.79(d^2h)^{0.60}$
		Crown	(11)	$Y_{crown}=6.92+181.84\ (d^2L)$	$\varepsilon=64.64(d^2L)^{0.95}$
	Ratio	Small branches	(12)	$R_1=0.08+e^{-0.26(Y_{crown})^{0.36}}$	$\varepsilon=0.12e^{-0.001(Ycrown)}$
		Coarse branches	(13)	$R_3=0.87Y_{crown}/(117.74+Y_{crown})$	$\varepsilon=0.12e^{-0.007(Ycrown)}$
		Medium branches	(14)	$R_2=1-(R_1+R_3)$	$\varepsilon=0.08e^{-0.001(Ycrown)}$
Cienciala et al. 2008	Biomass	Stem		$Y_{stem}=299.00(d)^{1.97}(h)^{0.98}$	$\varepsilon=218.82(d)^{1.47}$
		Stem wood		$Y_{wood}=237.02(d)^{2.00}(h)^{1.04}$	$\varepsilon=457.84(d)^{2.11}$
		Stem bark		$Y_{bark}=19.51(d)^{1.55}(h)^{0.96}$	$\varepsilon=79.80(d)^{1.91}$
		Crown		$Y_{crown}=34,378.88(d)^{3.06}(ht)^{-0.59}$	$\varepsilon=1,153.37(d)^{2.71}$
Suchomel et al. 2012	Biomass	Stem		$Y_{stem}=6,687.69(d)^{2.42}$	$\varepsilon=656.66(d)^{1.63}$
		Stem wood		$Y_{wood}=4,784.25(d)^{2.14}$	$\varepsilon=166.85(d)^{0.62}$
		Stem bark		$Y_{bark}=418.58(d)^{2.00}$	$\varepsilon=73.48(d)^{1.54}$
		Crown		$Y_{crown}=3,470.11(d)^{2.64}$	$\varepsilon=1,032.78(d)^{2.56}$

Y biomass of each studied component (kg), V stem volume (m^3), R_i biomass ratio of a specific branch fraction ($i=1$, small branches (sb); $i=2$, medium branches (mb); $i=3$, coarse branches (cb)) to the total crown biomass, *age* stand age (years), *d* over-bark diameter at breast height (m), *du* under-bark diameter at breast height (m), *ht* total height (m), *h* Delevoy height (m), *L* crown length (m)

[a] Model numbers refer to Table 3

effect) or to the β parameter (medium branches; multiplicative effect). The threshold stand ages around positive and negative intercept values were ca 51 and 21 years, for the total crown and small branch compartments, respectively. For the coarse branch component, the age effect was only captured through its influence on diameter and total height, as no covariate model could be fitted.

3.1.2 Validation data set

When applied to the validation data set, the statistical performances of the three empirical stem (total stem, stem wood, and stem bark) models were only slightly reduced compared to those obtained for the calibration data set in terms of both %RMSE and EF (Table 7). Although significant, the bias was acceptable. By contrast, the statistical performances of the four empirical crown models were strongly reduced compared to the calibration data set, leading to very low positive (total crown and coarse branches) or even negative (small and medium branches) EF values (Table 7). Figure 1 allows to compare the evolution of observed vs predicted biomass

values for these different components, as a function of diameter at breast height.

3.2 Structural approach

3.2.1 Calibration data set

As shown in Table 5, a common b_{stem} value of 0.52 was obtained for the stem volume model (Eq. 5). The modeling efficiency was very high (EF=0.99), and the relative root mean square error was 10.86 %. The *F*-test revealed an acceptable significant bias. The statistical performances obtained for the stem biomass compartments were quite similar to those observed for the corresponding components in the empirical approach (%RMSE 11.29–27.52; EF 0.89–0.99). Although the stem bark model was biased according to the *F*-test, visual inspection of the relationship between observed and estimated values still revealed no major bias (data not shown). The total (wood+bark) basic stem density obtained from the stem biomass model (parameter ρ_{stem}, Eq. 8) was estimated at 562 kg m^{-3} (Table 5). The basic wood density (Eq. 9, parameter $\rho_{stem\ wood}$) was estimated at 595 kg m^{-3}.

Table 6 Statistical performances of the selected biomass models on the calibration data set (the model parameters are given in Table 5)

		Bias			%RMSE	EF
		F-test[a]	Linear regression[b]			
		p value	Intercept	Slope		
Models of this study	Empirical approach					
	Stem	ns	−0.01 (−0.10, 0.07)	1.00 (0.99, 1.01)	11.69	0.99
	Stem wood	ns	−0.02 (−0.10, 0.06)	1.00 (0.99, 1.02)	12.86	0.98
	Stem bark	ns	−0.04 (−0.16, 0.09)	1.00 (0.97, 1.04)	25.73	0.90
	Crown	*	−0.60 (−0.94, -0.26)	1.14 (1.07, 1.21)	69.27	0.74
	Small branches	ns	0.10 (−0.35, 0.56)	0.95 (0.83, 1.07)	40.63	0.76
	Medium branches	*	−0.24 (−0.61, 0.13)	1.03 (0.92, 1.14)	48.30	0.73
	Coarse branches	*	−0.41 (−0.72, -0.09)	1.05 (0.97, 1.12)	60.14	0.87
	Structural approach					
	Stem	ns	−0.05 (−0.18, 0.08)	1.01 (0.98, 1.03)	11.29	0.99
	Stem wood	ns	−0.12 (−0.25, −0.00)	1.02 (1.00, 1.04)	11.60	0.99
	Stem bark	*	0.35 (0.20, 0.51)	0.91 (0.86, 0.95)	27.52	0.89
	Crown	ns	−0.09 (−0.38, 0.20)	1.00 (0.95, 1.06)	48.50	0.87
	Small branches	ns	0.37 (−0.10, 0.84)	0.90 (0.77, 1.03)	44.20	0.72
	Medium branches	*	−0.37 (−0.79, 0.05)	1.07 (0.94, 1.19)	52.94	0.68
	Coarse branches	ns	−1.03 (−1.33, −0.74)	1.17 (1.10, 1.24)	60.21	0.87
Models of literature	Cienciala et al. 2008					
	Stem	*	−0.17 (−0.30, −0.04)	1.03 (1.00, 1.05)	11.00	0.99
	Stem wood	ns	−0.12 (−0.25, −0.00)	1.02 (1.00, 1.04)	11.85	0.99
	Stem bark	ns	−0.03 (−0.19, 0.13)	1.00 (0.96, 1.05)	26.11	0.90
	Crown	ns	−0.04 (−0.31, 0.22)	1.00 (0.94, 1.05)	48.60	0.87
	Suchomel et al. 2012					
	Stem	*	−0.69 (−1.00, −0.38)	1.11 (1.05, 1.16)	28.59	0.92
	Stem wood	*	−1.82 (−2.18, −1.48)	1.28 (1.22, 1.34)	17.92	0.97
	Stem bark	*	−0.41 (−0.68, −0.15)	1.10 (1.02, 1.17)	41.91	0.75
	Crown	ns	−0.04 (−0.33, 0.24)	0.99 (0.94, 1.05)	47.26	0.88

ns not significant

*Significant at 5 %

[a] F-test performed on log-transformed values

[b] Linear regression between the log-transformed predicted and observed values

The total crown biomass model performed much better than the corresponding empirical model, in terms of both root mean square error (%RMSE=48.50) and modeling efficiency (EF=0.87); in addition, no significant bias was detected (Table 6). The selected structural model differed from the corresponding empirical one by using d^2L instead of d^2ht as the predictor (L=crown length (L=ht−h), d=diameter at breast height), as well as by using a fixed intercept value (α=6.92) instead of a simple linear age-dependent intercept. As a result, the age effect was only captured through its influence on diameter and crown length, without relying on any covariate. For the individual branch components, the precision of the biomass estimates was comparable (coarse branches) or slightly lower (medium and small branches) than for the empirical approach. Bias was either not detected or acceptable (medium branches). The parameters of the crown ratio models used for the biomass estimates of the corresponding branch categories are given in Table 5 and illustrated in Fig. 2. The small branch ratio model presented an EF of 0.71 with a relatively low root mean square error (%RMSE=27.17) but a significant bias. The coarse branch ratio model was characterized by an EF value of 0.82, a relatively low root mean square error (%RMSE=27.32) and a significant bias. The medium branch ratio model

Table 7 Statistical performances of the selected biomass models on the validation data set (the model parameters are given in Table 5)

		Bias			%RMSE	EF
		F-test[a]	Linear regression[b]			
		p-value	Intercept	Slope		
Models of this study	Empirical approach					
	Stem	*	−0.23 (−0.52, 0.06)	1.05 (1.00, 1.09)	16.34	0.93
	Stem wood	*	−0.25 (−0.55, 0.05)	1.05 (1.00, 1.10)	16.79	0.93
	Stem bark	*	−1.15 (−0.57, 0.28)	1.06 (0.95, 1.16)	21.95	0.84
	Crown	*	−0.22 (−1.15, 0.71)	1.17 (0.98, 1.36)	78.69	0.01
	Small branches	*	−0.32 (−1.04, 0.39)	1.26 (1.07, 1.45)	67.59	−0.17
	Medium branches	*	−0.90 (−2.52, 0.73)	1.42 (0.94, 1.89)	83.72	−0.31
	Coarse branches	ns	−0.06 (−0.92, 0.80)	1.06 (0.87, 1.24)	93.36	0.10
	Structural approach					
	Stem	*	0.11 (−0.14, 0.37)	0.99 (0.95, 1.03)	13.35	0.95
	Stem wood	*	−0.06 (−0.33, 0.21)	1.02 (0.98, 1.06)	14.27	0.95
	Stem bark	*	1.00 (0.65, 1.35)	0.78 (0.70, 0.87)	21.44	0.84
	Crown	*	−0.67 (−1.49, 0.14)	1.17 (1.02, 1.33)	48.12	0.63
	Small branches	*	−1.49 (−2.44, −0.54)	1.51 (1.27, 1.75)	63.20	−0.02
	Medium branches	*	−1.97 (−3.03, −0.90)	1.63 (1.34, 1.93)	65.75	0.19
	Coarse branches	ns	−0.47 (−1.10, 0.15)	1.11 (0.98, 1.24)	58.17	0.65
Models of literature	Cienciala et al. 2008					
	Stem	*	−0.00 (−0.27, 0.26)	1.01 (0.97, 1.05)	14.19	0.95
	Stem wood	*	−0.05 (−0.34, 0.24)	1.02 (0.97, 1.07)	15.09	0.94
	Stem bark	*	0.05 (−0.35, 0.45)	1.02 (0.92, 1.12)	24.28	0.80
	Crown	*	0.09 (−0.61, 0.80)	0.95 (0.82, 1.07)	67.11	0.28
	Suchomel et al. 2012					
	Stem	*	0.98 (0.63, 1.34)	0.84 (0.79, 0.90)	27.86	0.80
	Stem wood	*	0.05 (−0.33, 0.43)	0.98 (0.92, 1.04)	18.56	0.91
	Stem bark	*	0.88 (0.38, 1.38)	0.79 (0.67, 0.91)	30.49	0.69
	Crown	ns	−0.44 (−1.23, 0.34)	1.06 (0.92, 1.20)	44.23	0.69

The relationships between observed or predicted biomass by the four models and diameter at breast height are depicted in Fig. 1 for the different tree components

ns not significant

*Significant at 5 %

[a] F-test performed on log-transformed values

[b] Linear regression between the log-transformed predicted and observed values

showed a relative root mean square error of 31.18 % and an EF of 0.48 with a nonsignificant bias.

3.2.2 Validation data set

The evolution of observed and predicted biomass for the validation data set is depicted in Fig. 1 for each compartment. The application of the stem structural models to the validation data set only resulted in slightly reduced statistical performances in terms of precision (small EF decrease for all models, slight increase in %RMSE for the total stem and stem wood models). The bias appeared to be acceptable for the total stem and stem wood models; for the stem bark compartment, the bias, already detected on the calibration data set, was still present. By contrast to the stem models, the decrease in statistical performance resulting from fitting the models on the validation data set was much more pronounced for the structural crown models, especially for the small and medium branches. Compared to the corresponding empirical models, the decrease of the modeling efficiencies for the structural total crown and coarse branches models was however much more limited. For those two components, EFs were about 0.63–0.65 and the bias was either not detected (coarse branches) or very small.
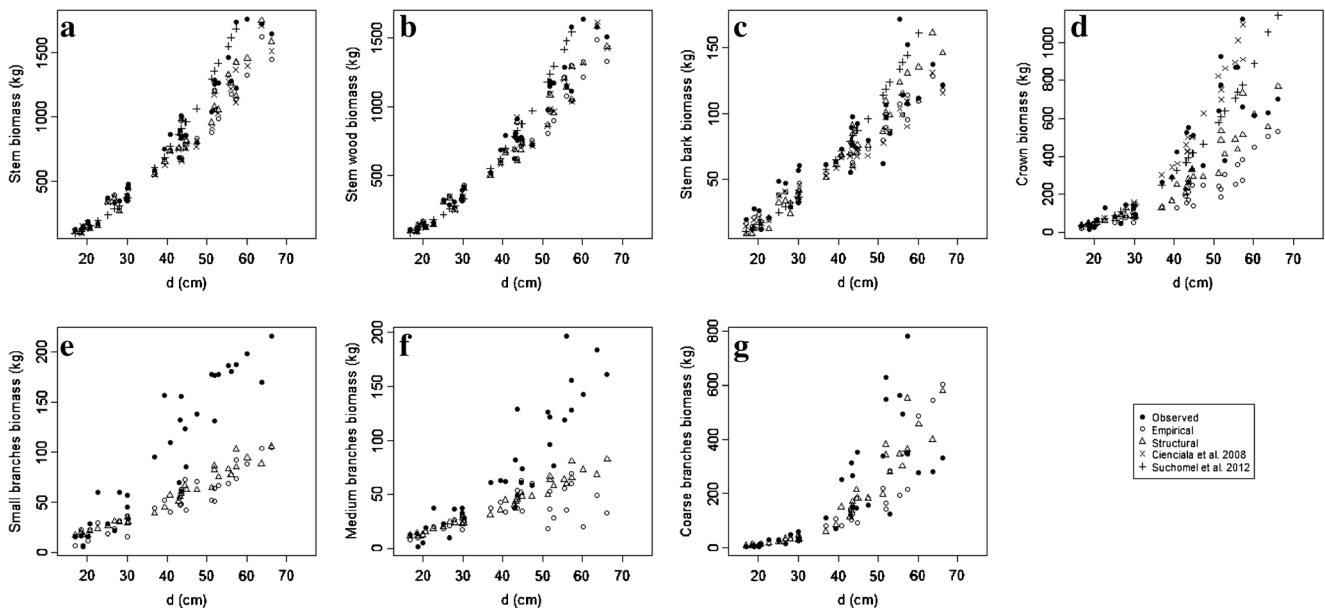
Fig. 1 Observed and predicted biomass (kg) as a function of diameter at breast height (d (cm)) for the validation data set. **a** Total stem, **b** stem wood, **c** stem bark, **d** total crown, **e** small branches, **f** medium branches, and **g** coarse branches. The coefficients of the models are presented in Table 5, and their statistical performances on the validation data set are given in Table 7

4 Discussion

4.1 Empirical vs structural approach in the calibration data set

4.1.1 Empirical approach

The reference equation we used in the empirical approach (Eq. 1) integrates both diameter at breast height and height (ht or h, depending on the component). The parameter α represents the biomass of the tree before it attains a height of 1.30 m, the parameter β represents the scaling factor, and the parameter γ represents the allometric exponent governing the proportionality between biomass and volume increments (Genet et al. 2011). It was selected as it proved to be quite efficient to model the different biomass compartments of beech (*Fagus sylvatica* L.) over a wide range of stands located in Central Europe (Genet et al. 2011). Sessile and pedunculate oaks are indeed expected to behave similarly to beech in terms of allometric relationships (André et al. 2010).

For the stem and stem wood components, all base models performed quite similarly in terms of BIC and AIC. The local fitting of the parameter(s) only slightly increased the model performances, and the differences between models remained limited. In addition, this step resulted in greatly increasing total number of parameters of the models, as well as the correlation between them (data not shown). For the other components, one (stem bark and small branches, model 4; total crown, model 3) or two (medium and coarse branches, models 3 and 4) global base models were found to have much lower performances. For the total crown, the local fitting of the

worst global base model (model 3) surprisingly resulted in obtaining the best local model (model 3.2); because of quite similar AIC and BIC values, the latter could only be selected over model 1.5 after performing an *LRT*. Overall, these results suggest that using purely statistical criteria could easily change both the form and the parametrization of the models.

With the exception of the coarse branches, it was possible to find a suitable covariate model for all components. This allowed to greatly reduce the total number of parameters to be estimated without significantly affecting the performances of the models. In addition, while the local fitting would only apply to the corresponding stands, a major advantage of the covariate models is that they can possibly be applied to a large range of stands provided the value of the covariate is known for the stands in question. In the present case, we could only use stand age as a covariate. Interpretation of this so-called age effect is however difficult for at least three reasons: (i) part of the age effect is already included in the values of the predictors; (ii) due to the relatively limited number of stands, the so-called age effect cannot be separated from other factors which are also expected to change with age such as stand density; (iii) it is tightly model-dependent, especially because the selection of a specific model form interacts with parameter estimates. The age effect was thoroughly discussed in Genet et al. (2011).

4.1.2 Structural approach

The performance of the so-called constant form volume model (Spurr 1952; Husch et al. 1982) we used as starting point of

INRA SCIENCE & IMPACT   Springer

the structural approach was compared to that of the three other global models used for modeling the biomasses in the empirical approach (models 1, 2, and 3 in Table 3 with $Y=V$); based on the LRT, the constant form volume model proved to be the most appropriate (data not shown). Using this model, the product of the estimated b_{stem} parameter value (0.52) times the constant $(4/\pi)$, i.e., 0.66, can be interpreted as the ratio of the stem volume to that of a cylinder having the same diameter at breast height and the same height (form factor). This value (0.66) was not significantly different from that (0.62) obtained by the model of Vallet et al. (2006) (p value of the paired t test =0.06). The failure to adjust a covariate model to the local estimates of the b_{stem} parameter suggests that stand age was not governing the form factor. The combined (wood+bark) density estimate of 560 kg m^{-3} obtained from the stem biomass model (Eq. 8) compared quite well with published values (468–595 kg m^{-3}) (Deret-Varcin 1983; Nepveu 1984). Although this stem model differed from the corresponding empirical one by its basic form and its global instead of local/covariate fitting, it had quite similar statistical performances. This again suggests some caution when interpreting the values of the parameters obtained by using a purely statistical approach. The structural stem bark model was quite similar to the empirical model in terms of %RMSE and EF; it however presented a limited but significant bias. With the structural approach, it was necessary to have an estimate of both total stem and stem wood biomasses for the assessment of the stem bark component. As we did not measure the under-bark diameter, we had to predict it from an existing equation; in addition, we also had to hypothesize the similarity of the under- and over-bark form factors. To test for the reliability of the under-bark diameter values obtained from applying the equation of Dagnelie et al. (1999), we compared our estimates to those obtained using a simple ratio of 0.926 between under- and outside-bark diameter (Dhôte et al. 2000); although no significant differences were observed between both values (p value of the paired t test =0.08), we could still not discard a possible bias in our under-bark diameter estimates. By contrast, we could clearly demonstrate the similarity of the outside-bark and under-bark form factors, using an independent data set of 963 oak trees for which both under- and outside-bark volumes had been measured (p value of the paired t test =0.62, data not shown). Despite the abovementioned limitations, the estimate of the basic stem wood density (595 kg m^{-3}, Eq. 9) was in close agreement with published value of 580 kg m^{-3} (IPCC 2003).

In the structural approach, the statistical performances of the models estimating the various branch components tightly rely on the total crown model as well as on the ratio models. By using L as an estimate of crown size instead of total height, the structural approach was much more successful than the empirical one in estimating the total crown biomass. In the present case, however, L was only a proxy of crown length as

it corresponds to the difference between total height and the height at which stem diameter is half the diameter at 1.3 m height. Although its proximity to the effective crown length is expected to depend on silviculture in general and on stand density in particular, this empirical height proved quite useful in estimating stem wood volumes for management purposes in broad-leaved stands (André 2007); in addition, it can easily be calculated from standard biomass data, without any additional field height measurements. The ratio approach was first developed by Honer (1964) to estimate merchantable timber volume at different cutting sections. Since then, it has also been widely used in biomass studies, to derive estimates for components that are difficult to assess such as foliage or roots from more easily predicted ones (Parresol et al. 1987; Bartelink 1996) and rarely used for the estimation of branch category biomass (Seifert et al. 2006). Because the ratios of the small and coarse branches followed a characteristic pattern when plotted against the total crown biomass (Fig. 2), we decided to first model the proportions of those two subcompartments and to obtain the ratio of the medium branches as the difference between one and the sum of the two other branch ratios; this approach allowed us to cope with the large variability observed for this branch component, while ensuring the additivity. In the small branch ratio model, θ_0 represents the asymptotic proportion of small branches at increasingly higher total crown biomasses (ca 8 %), whereas both θ_1 and θ_2 determine the rate of decrease of the small branch proportion with an increase in total crown biomass. More specifically, θ_1 is the rate of decrease in the small branches proportion for each unit increase in total crown biomass; θ_2 is the mean amplitude $(R_{1max}-\theta_0)/2)$ of the small branch ratio. For the coarse branch model, θ_3 represents the asymptotic proportion of coarse branches or upper limit obtained at large total crown biomasses (ca 87 %), while θ_4 is the total crown biomass obtained when the ratio of the coarse branches reaches half of its upper limit $(\theta_3/2)$.

4.2 Explaining the contrasting model performances on the validation data set

To further test both approaches, we compared the statistical performances of our models and of two other oak biomass models available in the literature (Cienciala et al. 2008; Suchomel et al. 2012) on an independent data set of 33 sessile oak trees from two forest stands (Table 7, Fig. 1). These stands were purposely selected because they included silvicultures well outside the range used in the calibration data set, allowing to test for the robustness of the approaches.

For the total stem and stem wood components, the models containing both d and h as predictors performed better than the simple diameter models of Suchomel et al. (2012) when fitted based on the calibration data set. In addition, their modeling
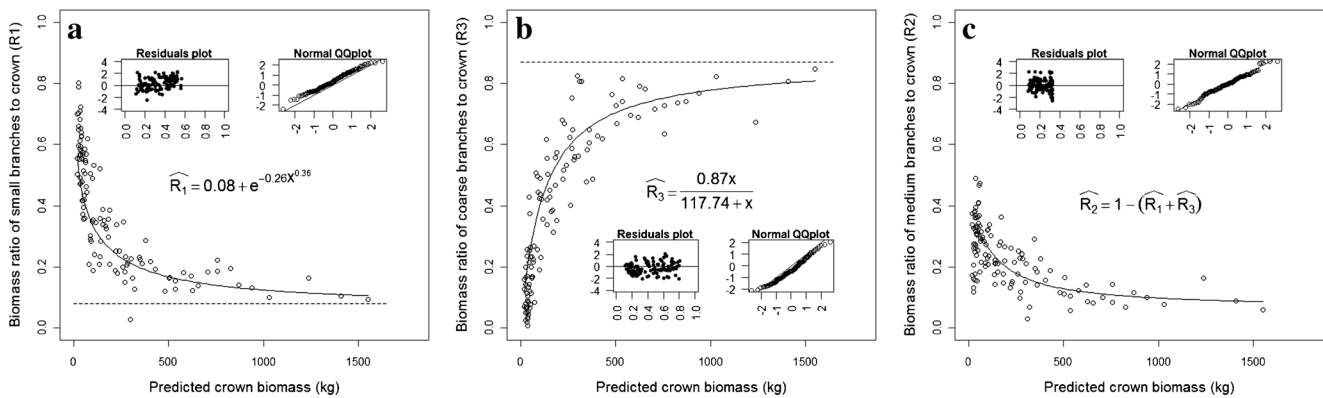
Fig. 2 Biomass ratio of small (**a**), coarse (**b**), and medium (**c**) branches to total crown, as a function of total crown predicted biomass (kg) for the calibration data set. A *point* represents a measured ratio, the *solid line* represents the model fit, and the *dotted line* represents the asymptote. The statistical performances of each model are given in the text

efficiencies were only slightly reduced when shifting to the validation data set. This, together with the failure to adjust any covariate model for the parameters used in the structural approach, suggests that *d* and *h* captured much of the variability associated with stem biomass, which is in agreement with many other studies (e.g., Cairns et al. 2003; Chave et al. 2004; Brandeis et al. 2006; António et al. 2007). Quite interestingly, the multiplicative model of Cienciala et al. (2008) leads to estimates of parameters for *d* and *h* that were quite similar to those imposed by the structural approach, which strongly supports the latter.

For both the calibration and validation data sets, the stem bark biomass was equally well predicted by all three models that included *d* and *h* as predictors, despite strong differences in model forms. Also, the decline in the statistical performances when shifting from the calibration to the validation data step remained relatively limited. As for the stem and stem wood components, the further inclusion of height as a predictor in stem bark models increased their predictive ability, but the way height contributed to these increased performances remained unclear in both the empirical and multiplicative approaches. In particular, the exponents estimated for *d* and *h* in the model of Cienciala et al. (2008) cannot readily be explained. In this respect, the major advantage of the structural approach resides in the clear interpretation of its parameters, for otherwise similar statistical performance.

For the whole crown level, it appeared that the structural approach, the multiplicative *d* and *ht* model, and the simple allometric *d* equation resulted in similar and satisfactory statistical performances considering the calibration data set; the empirical crown biomass model that included both *d* and *ht* was much less performant. This suggests that, contrary to the stem components, inclusion of *ht* does not systematically translate into higher modeling performance for crown. This was largely confirmed when looking at the results of the independent validation data set. The degree of deterioration

in statistical performances was totally independent on whether the model included *d* and *ht*, or only *d*. For instance, the empirical total crown model and the multiplicative model of Cienciala et al. (2008) were found to be inefficient, while the two other models performed similarly. Possible explanations of these contrasting patterns are to be found in the specific model forms. When using d^2ht as the predictor in the empirical approach, similar stem volumes could be obtained for quite contrasting values of *d* and *h*. For a given stem volume, however, one expects the crown biomass to be more important in reduced competition conditions, that is for relatively lower values of the *ht/d* ratio. On the other hand, increasing total stem biomass does not necessarily translate into increased crown biomass due to the accumulating pattern of stem increment. In the empirical approach, the required correcting factor is given by the age-dependent additive intercept component that increases or decreases the biomass depending on stand age (cf. 3.1.1). With the multiplicative *h* and *d* model, the abovementioned competition and ontogeny effects are partly accounted for through the exponents associated to *d* and *h*. Considering the exponent values imposed by the empirical approach (2 and 1 for *d* and *h*, respectively), the parameter estimates obtained for *d* and *h* in the multiplicative model can be interpreted as if this model was multiplied by a factor (d^γ/h^δ). In this respect, the crown biomass is adjusted by a function accounting for competition. In the present case, the diameter coefficient was much greater than 2 (3.06), while the height coefficient was negative (−0.59). Because of the negative exponent associated with *h*, trees of given diameters tend to have increasingly smaller crown biomass with increase in total height. Although competition seemed to be more accounted for in that approach compared to the empirical models, application to an independent data resulted in very poor crown biomass estimates. A possible reason for this failure could be the difference in competition conditions between the two data sets. The success of the simple *d* model in estimating the crown biomass for both data sets was rather

unexpected, though the much greater exponent value associated with d compared to h in the model of Cienciala et al. (2008) suggests a prominent role of d over h. The same conclusion was drawn by Lambert et al. (2005) in a systematic comparison of only dbh-, and dbh- and height-based set of equations.

The major difference between the structural approach and the previous ones is that in addition to the diameter of some stem section, the equation incorporates a more direct estimate of the crown size through crown length (L) (Carvalho and Parresol 2003; Duursma et al. 2007; Cienciala et al. 2008). This property proved to be quite efficient in limiting the loss of statistical performance when shifting to the validation data set. By contrast, it fails to adequately predict the biomass of the different branch categories, with the notable exception of the coarse branches.

5 Conclusions and perspectives

In this study, we explored a number of possible alternatives to increase the robustness and genericity of biomass equations, based on a large sample of 150 oak trees (*Q. robur*, *Q. petraea*) spanning a wide range of site conditions and stand characteristics. These alternatives were tested using up-to-date statistical methods taking into account the heteroscedasticity of the residuals and the additivity constraint.

A first alternative is to rely on covariate models that allow to partly explain the variation in parameter values by some site or stand characteristics and stand age in our case. The contrasting changes in statistical performances observed for the stem vs the crown models when shifting from the calibration to the validation data sets in the empirical approach clearly showed that the success of the covariate models largely depended on the selection of an appropriate reference model. In this respect, one possible drawback of the covariate approach is that, because of its flexibility, it tends to give satisfactory statistical fitting for a much larger set of candidate models, including models with quite contrasting forms and predictor(s). The results obtained for the total stem and stem wood components also showed that this type of model can result in overparametrization when priority is given to fitting (empirical approach) over interpretation (structural approach) in model selection.

A second way to deal with robustness and genericity is through the use of a multiplicative function of d and h, where the exponents are parameters to be estimated by the model. When used at the stem level, the comparison between that approach and the structural one in which the exponents associated to d and h were fixed a priori to 2 and 1, respectively, showed that it could result in overparametrization of the model. Applying this procedure at the crown level resulted in unsatisfactory statistical performances for the independent

validation data set despite excellent results on the calibration data set. As a result, great caution is recommended in using that approach outside the calibration range.

The third alternative we tested to improve the robustness and genericity of the biomass equations is through the development of biologically and physically meaningful models. For each of the component, it first implied to identify the relevant set of predictors and the way to combine them. In this respect, our results clearly showed the importance of selecting specific sets of predictors for the stem vs the crown components, instead of using comparable ones (d^2h or d^2ht) as was done in the empirical approach. Using structural models proved to be quite successful in predicting both the stem components and the total crown biomass even for the independent data set. Beyond the performance in terms of robustness and genericity, additional advantages of this approach are the reduction in the total number of parameters to be estimated (parsimony) as well as the possibility to interpret the models.

Capitalizing on our results, several avenues can now be identified to further improve the structural approach. At the stem level, a validated under-bark diameter model would possibly still improve the model performances in terms of %RMSE, modeling efficiency, and bias for the stem bark component. In the absence of such a model, a recommendation would be to measure both under- and over-bark diameters while carrying out biomass sampling. At the crown level, our suggestion is to test for a larger set of crown size measurements to serve as possible alternatives for L while simultaneously refining our branch ratio models.

References

Analla M (1998) Model validation through the linear regression fit to actual versus predicted values. Agric Syst 57:115–119. doi:10.1016/S0308-521X(97)00073-5

André F (2007) Influence of the heterogeneity of canopy structure on the spatio-temporal variability of atmospheric deposition within a mixed oak-beech stand. PhD thesis, Université catholique de Louvain

André F, Ponette Q (2003) Comparison of biomass and nutrient content between oak (*Quercus petraea*) and hornbeam (*Carpinus betulus*)

trees in a coppice-with-standards stand in Chimay (Belgium). Ann For Sci 60:489–502

André F, Jonard M, Ponette Q (2010) Biomass and nutrient content of sessile oak (*Quercus petraea* (Matt.) Liebl.) and beech (*Fagus sylvatica* L.) stem and branches in a mixed stand in southern Belgium. Sci Total Environ 408:2285–2294. doi:10.1016/j.scitotenv.2010.02.040

António N, Tomé M, Tomé J, Soares P, Fontes L (2007) Effect of tree, stand, and site variables on the allometry of *Eucalyptus globulus* tree biomass. Can J For Res 37:895–906. doi:10.1139/X06-276

Arevalo C, Volk TA, Bevilacqua E, Abrahamson L (2007) Development and validation of aboveground biomass estimations for four Salix clones in central New York. Biomass Bioenergy 31:1–12

Bartelink H (1996) Allometric relationships on biomass and needle area of Douglas-fir. For Ecol Manag 86:193–203

Bi H, Turner J, Lambert MJ (2004) Additive biomass equations for native eucalypt forest trees of temperate Australia. Trees 18:467–479. doi:10.1007/s00468-004-0333-z

Brandeis TJ, Delaney M, Parresol BR, Royer L (2006) Development of equations for predicting Puerto Rican subtropical dry forest biomass and volume. For Ecol Manag 233:133–142. doi:10.1016/j.foreco.2006.06.012

Cairns MA, Olmsted I, Granados J, Argaez J (2003) Composition and aboveground tree biomass of a dry semi-evergreen forest on Mexico's Yucatan Peninsula. For Ecol Manag 186:125–132. doi:10.1016/S0378-1127(03)00229-9

Canadell JG, Raupach MR (2008) Managing forests for climate change mitigation. Science 320:1456–1457. doi:10.1126/science.1155458

Carroll RJ, Wu JCF, Ruppert D (1988) The effect of estimating weights in weighted least squares. J Am Stat Assoc 83:1045–1054. doi:10.2307/2290134

Carvalho JP, Parresol BR (2003) Additivity in tree biomass components of Pyrenean oak (*Quercus pyrenaica* Willd.). For Ecol Manag 179:269–276. doi:10.1016/s0378-1127(02)00549-2

Chave J, Condit R, Aguilar S, Hernandez A, Lao S, Perez R (2004) Error propagation and scaling for tropical forest biomass estimates. Philos Trans R Soc Lond Ser B Biol Sci 359:409–420. doi:10.1098/rstb.2003.1425

Cienciala E, Apltauer J, Exnerová Z, Tatarinov F (2008) Biomass functions applicable to oak trees grown in Central-European forestry. J For Sci 54:109–120

Dagnelie P (1975) Analyse statistique à plusieurs variables. Presses Agronomiques de Gembloux, Gembloux

Dagnelie P, Palm R, Rondeux J (1999) Tables de cubage des arbres et des peuplements forestiers. Presses Agronomiques de Gembloux, Gembloux

Dent JB, Blackie MJ (1979) Systems simulation in agriculture. Applied Science Publishers Ltd., London

Deret-Varcin E (1983) Etude comparative de la qualité du bois de trois types de chênes (rouvres, pédonculés et intermédiaires), en forêt de Morimond. Ann For Sci 40:373–398

Dhôte J-F, Hatsch E, Rittié D (2000) Forme de la tige, tarifs de cubage et ventilation de la production en volume chez le Chêne sessile. Ann For Sci 57:121–142

Duursma R, Marshall J, Robinson A, Pangle R (2007) Description and test of a simple process-based model of forest growth for mixed-species stands. Ecol Model 203:297–311. doi:10.1016/j.ecolmodel.2006.11.032

FAO (2010) Global Forest Resources Assessment 2010—main report. FAO Forestry Paper No. 163. http://www.fao.org/docrep/013/i1757e/i1757e00.htm. Accessed 12 June 2013

Fortin M, Ningre F, Robert N, Mothe F (2012) Quantifying the impact of forest management on the carbon balance of the forest-wood product chain: a case study applied to even-aged oak stands in France. For Ecol Manag 279:176–188. doi:10.1016/j.foreco.2012.05.031

Genet A, Wernsdörfer H, Jonard M, Pretzsch H, Rauch M, Ponette Q, Nys C, Legout A, Ranger J, Vallet P, Saint-André L (2011) Ontogeny partly explains the apparent heterogeneity of published biomass equations for *Fagus sylvatica* in central Europe. For Ecol Manag 261:1188–1202. doi:10.1016/j.foreco.2010.12.034

Honer T (1964) The use of height and squared diameter ratios for the estimation of merchantable cubic foot volume. For Chron 40:324–331

Huet S, Bouvier A, Poursat M-A, Jolivet E (2004) Statistical tools for nonlinear regression: a pratical guide with S-PLUS and R examples. Springer, New York

Husch B, Miller C, Beers T (1982) Forest mensuration. Wiley, New York

IPCC (2003) Good practice guidance for land use, land-use change and forestry. Institute for Global Environmental Strategies, Hayama

Ketterings QM, Coe R, van Noordwijk M, Ambagau Y, Palm CA (2001) Reducing uncertainty in the use of allometric biomass equations for predicting above-ground tree biomass in mixed secondary forests. For Ecol Manag 146:199–209. doi:10.1016/s0378-1127(00)00460-6

Lambert M, Ung C, Raulier F (2005) Canadian national tree aboveground biomass equations. Can J For Res 35:1996–2018. doi:10.1139/X05-112

Lehtonen A, Mäkipää R, Heikkinen J, Sievänen R, Liski J (2004) Biomass expansion factors (BEFs) for Scots pine, Norway spruce and birch according to stand age for boreal forests. For Ecol Manag 188:211–224. doi:10.1016/j.foreco.2003.07.008

Loague K, Green RE (1991) Statistical and graphical methods for evaluating solute transport models: overview and application. J Contam Hydrol 7:51–73. doi:10.1016/0169-7722(91)90038-3

Longuetaud F, Santenoise P, Mothe F, Senga Kiessé T, Rivoire M, Saint-André L, Ognouabi N, Deleuze C (2013) Modeling volume expansion factors for temperate tree species in France. For Ecol Manag 292:111–121. doi:10.1016/j.foreco.2012.12.023

Mayer D, Butler D (1993) Statistical validation. Ecol Model 68:21–32

Mayer D, Stuart M, Swain A (1994) Regression of real-world data on model output: an appropriate overall test of validity. Agric Syst 45:93–104

McKechnie J, Colombo S, Chen J, Mabee W, MacLean HL (2010) Forest bioenergy or forest carbon? Assessing trade-offs in greenhouse gas mitigation with wood-based fuels. Environ Sci Technol 45:789–795

Návar J, Nájera J, Jurado E (2002) Biomass estimation equations in the Tamaulipan thornscrub of north-eastern Mexico. J Arid Environ 52:167–179. doi:10.1006/jare.2001.0819

Nelson BW, Mesquita R, Pereira JL, Aquino G, de Souza S, Teixeira Batista G, Bovino Couto L (1999) Allometric regressions for improved estimate of secondary forest biomass in the central Amazon. For Ecol Manag 117:149–167. doi:10.1016/s0378-1127(98)00475-7

Nepveu G (1984) Contrôle héréditaire de la densité et de la rétractibilité du bois de trois espèces de Chêne (*Quercus petraea, Quercus robur et Quercus rubra*). Silvae Genet 33:110–115

Parresol BR (1999) Assessing tree and stand biomass: a review with examples and critical comparisons. For Sci 45:573–593

Parresol BR (2001) Additivity of nonlinear biomass equations. Can J For Res 31:865–878. doi:10.1139/cjfr-31-5-865

Parresol BR, Hotvedt JE, Cao QV (1987) A volume and taper prediction system for bald cypress. Can J For Res 17:250–259

Pinheiro JC, Bates DM (2000) Mixed-effects models in S and S-PLUS. Springer, New York

Prisley SP, Mortimer MJ (2004) A synthesis of literature on evaluation of models for policy applications, with implications for forest carbon accounting. For Ecol Manag 198:89–103. doi:10.1016/j.foreco.2004.03.038

Ranger J, Marques R, Colin-Belgrand M, Flammang N, Gelhaye D (1995) The dynamics of biomass and nutrient accumulation in a Douglas-fir (*Pseudotsuga menziesii* Franco) stand studied using a chronosequence approach. For Ecol Manag 72:167–183. doi:10.1016/0378-1127(94)03469-d

Sabine CL, Heimann M, Artaxo P, Bakker DC, Chen C-TA, Field CB, Gruber N, Le Quéré C, Prinn RG, Richey J (2004) Current status

and past trends of the global carbon cycle. Scope-Sci Commitee Probl Environ Int Counc Sci Unions 62:17–44

Saint-André L, M'Bou AT, Mabiala A, Mouvondy W, Jourdan C, Roupsard O, Deleporte P, Hamel O, Nouvellon Y (2005) Age-related equations for above- and below-ground biomass of a Eucalyptus hybrid in Congo. For Ecol Manag 205:199–214. doi:10.1016/j.foreco.2004.10.006

Sakamoto Y, Ishiguro M, Kitagawa G (1986) Akaike information crite-rion statistics. Reidel, Dordrecht, Holland

SAS Institute Inc (2011) SAS/ETS OnlineDoc® 9.3. SAS Institute Inc, Cary, NC

Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6:461–464

Seifert T, Schuck J, Block J, Pretzsch H (2006) Simulation von Biomasse- und Nährstoffgehalt von Waldbäumen. Beiträge zur Jahrestagung vom 29:208–223

Spurr SH (1952) Forest inventory. Ronald Press, New York

Suchomel C, Pyttel P, Becker G, Bauhus J (2012) Biomass equations for sessile oak (*Quercus petraea* (Matt.) Liebl.) and hornbeam (*Carpinus betulus* L.) in aged coppiced forests in southwest Germany. Biomass Bioenergy 46:722–730. doi:10.1016/j.biombioe.2012.06.021

Ter-Mikaelian MT, Korzukhin MD (1997) Biomass equations for sixty-five North American tree species. For Ecol Manag 97:1–24. doi:10.1016/s0378-1127(97)00019-4

Thornton P, Hansen J (1996) A note on regressing real-world data on model output. Agric Syst 50:411–414

Vallet P, Dhôte J-F, Le Moguédec G, Ravart M, Pignard G (2006) Development of total aboveground volume equations for sev-en important forest tree species in France. For Ecol Manag 229:98–110

Vanclay JK, Skovsgaard JP (1997) Evaluating forest growth models. Ecol Model 98:1–12

Wutzler T, Wirth C, Schumacher J (2008) Generic biomass functions for Common beech (*Fagus sylvatica*) in Central Europe: predictions and components of uncertainty. Can J For Res 38:1661–1675. doi:10.1139/x07-194

Zanetto A, Roussel G, Kremer A (1994) Geographic variation of inter-specific differentiation between *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. For Genet 1:111–123